Retrieval, Compilation, and Exploratory Analyses of Voucher Specimens for Vertebrates and Vascular Plants in 14 Northeastern National Parks

Allan O'Connell & Andrew Gilbert
USGS, Patuxent Wildlife Research Center
Laurel, MD

Photo by A. O'Connell

Table of Contents

Appendix 1.   Database of institutions with natural history collections searched and contacted. The information is in MS-Access format and is located at Patuxent headquarters in Laurel, MD and the NPS Inventory and Monitoring Program at the University of Rhode Island, Kingston.

Appendix 2.   Records of all voucher specimens collected from museums and herbaria originating within or near 9 northeast temperate and 5 coastal parks. This information is in MS-Excel format and is located at Patuxent headquarters in Laurel, MD and and the NPS Inventory and Monitoring Program at the University of Rhode Island, Kingston.

List of Tables (located at end of document)

List of Figures (located at end of document)

Executive Summary

As part of the National Park Service (NPS) effort to implement biological inventories and long-term monitoring of park resources, we searched natural history collections to determine the number of vertebrate (mammals, birds, amphibians, reptiles) and vascular plant voucher specimens originating within or near select national park unit boundaries throughout the northeastern United States. These records provided a measure of historical biodiversity for the region with respect to individual park units. We queried 274 collection managers curating 307 natural history collections and received a 70% response rate. Searches of collections placed specimens into one of four locality categories, ranging from within current park boundaries to a county within the state where the park was located. We evaluated park characteristics (e.g., age, size) to evaluate their effect on the number of specimens found. We assembled >30,000 specimen records including over 4,000 (21%) determined to originate within current park boundaries. Distribution of specimens was fairly uniform between 1890 and 2000 with declines in collecting corresponding roughly to world events (i.e., wars, great depression). Manual searches showed that specimens could be found in collections of varying sizes but require considerable time and effort. Because most collections are not yet computerized, manual searches remain the only viable option if information is needed for immediate use.. We found specimens originating within or near the boundaries of parks ranging from small collections in local educational institutions to the world's largest museums, some far outside the region of origin, as far away as Europe. Thus, global searches will be necessary to locate specimens of interest. We discuss how to conduct efficient searches of natural history collections, the value of voucher specimens in documenting local and regional biodiversity, and the utility of estimating species richness.

Introduction

      Most of the scientific data at our disposal are in the form of museum and herbaria collections and their importance lies in the transfer of information about biodiversity (Burgman et al. 1995). Scientific collections validate the presence of a species from specific localities and serve as an historical reference for the geographic distribution of biodiversity (Fagen and Kareiva 1997). Further, natural history collections contain data critical to decisions about biodiversity conservation (Ponder 1999) and serve as a basis for taxonomic reference and information (Hawksworth and Mound 1991). With the threat of extinction increasing, the importance of voucher specimens cannot be understated. Nevertheless, there seems to be a widespread disregard for the fundamental value of natural history collections that threatens the integrity of biological knowledge (Cotterill 1995). Collections are often ignored because records are often difficult to find or access (Cotterill 1997), and as a result, the initial appeal and application of the data are limited. Furthermore, critics have questioned the usefulness of museum and herbaria collections due to a lack of replication in sampling strategies (Ponder et al 2001), variability in collection effort (McCarthy 1998), loose compilation criteria (Alberch 1993), poor curatorial preservation procedures, and lack of coherent administrative policies for collecting efforts.

      Although efforts like the National Biological Information Infrastructure (NBII) and the Electronic Natural History Initiative are encouraging, use of data from scientific collections, particularly voucher specimens, remain an enigma to many in terms of their value and utility. The problem becomes even more vexing when dealing with multiple taxa across physiographic regions or political boundaries. To date, most attempts at using this information have examined data focusing on specific taxa with unique characteristcis (Ponder et al. 2001), correcting for

collection effort (Fagen and Kareiva 1997), mapping species distributions or determining areas of conservation importance (Ponder et al. 2001). A critical, often overlooked aspect of using this type of information is to develop an historical framework from which to compare records for future monitoring of biodiversity that can both document and track temporal changes. An important component of using this information is to first determine the full complement of records available. For example, in the use of butterfly lists to make biodiversity comparisons in Oregon (Fagen and Kareiva 1997), < 50% of the dataset was used because the remainder of the dataset was not computerized.

The NPS has recently implemented a system-wide inventory and monitoring program to assess biodiversity throughout a system of physiographic networks (NPS 2000). As part of this effort, some networks have attempted to locate voucher specimens originating on NPS lands thereby attempting to build a database of historical biodiversity. Recommendations from previous work include searching for records of a single park at a time, using subject matter experts to focus on familiar taxa, and conducting manual searches of collections (Bennett 2001). Considering that collection of natural history specimens predates the establishment of many parks, compilation of voucher specimen records will require a combination of organization, innovation, and trial and error.

Our primary objective was to locate specimens originating within or near national park boundaries of the Northeast Temperate Network: Acadia National Park (ACAD), Marsh-Billings-Rockefeller National Historical Park (MABI), Minute Man National Historical Park (MIMA), Morristown National Historical Park (MORR), Roosevelt-Vanderbilt National Historic Site (ROVA), Saint-Gaudens National Historic Site (SAGA), Saugus Iron Works National Historic Site (SAIR), Saratoga National Historical Park (SARA), and Weir Farm National

Historic Site (WEFA); and the Coastal and Barrier Island Network: Assateague Island National

Seashore (ASIS), Cape Cod National Seashore (CACO), Fire Island National Seashore (FIIS),

Gateway National Recreation Area (GATE), and Sagamore Hill National Historic Site (SAHI)

(Table 1). Secondary objectives included evaluating the importance of park characteristics (e.g.,

size, age) in generating specimen records, the effectiveness of manually searching different size

and type collections and the value of using voucher specimens to estimate species richness.

Methods

Data requests

We obtained information about vertebrate (except fish) and vascular plant natural history

collections by first searching two web-accessible databases of natural history collections: the

Index Herbariorum (IH) (http://www.nybg.org/bsci/ih/ih.html) and the Directory of Research

Systematics Collections (DRSC) (http://www.nbii.gov/datainfo/syscollect/drsc/). We also sent

out requests for information about collections to several e-mail list-serves (TWS-L, NHCOLL-L,

ORNITH-L) and obtained a list of museum contacts from John Karish (NPS, Philadelphia

Support Office) from a similar project. Additional collection information was found by

searching websites of regional biology departments.

We mailed requests for data to 274 collection managers curating 299 natural history

collections and 8 state natural heritage programs. We specifically requested data for specimens

originating within the 14 northeastern national parks. Information about natural history

collections were recorded in a Microsoft Access 2000 database. Collection information were

separated by taxa (e.g. Cornell University Museum of Vertebrates ornithology collection) where

taxa-specific data were available. Information such as size of collection, percentage

computerized, contact person and address, web address, and notes about the collections were recorded. We determined that much of the information provided in the two natural history collection databases were out of date; therefore, we checked contact information for all institutions through websites or by contacting institutions directly and updated information as necessary.

To reduce search time and increase the number of responses from institutions, we broadened search criteria to county-wide locality requests. This approach also had the benefit of including locations that were miss-spelled or used historic names, which may have otherwise excluded this information. We sent institutions a list of parks and localities by state and county(s). We requested that the following data fields be provided: park name, taxonomic name, common name, catalog number, accession number, condition of specimen, collector's name, date of collection, locality information, latitude-longitude, and comments. We e-mailed follow up requests for data to 177 collection managers who did not respond within 6 weeks of the initial request for data. We logged responses into the collection database as they were received.

Searches of Specimen Collections

We searched 22 collections, 12 via the Internet and 10 manually. Those collections that we searched manually were selected from the pool of collections that were not (or only partially) computerized and from which we had not received data (i.e., response to our mailing). We then randomly selected five vertebrate and five plant collections from each of three size categories: small (≤30,000), medium (≤195,000), and large (>195,000) and subsequently chose at least one specimen collection from each of these categories. Size categories were determined by dividing the size distribution of collections in thirds. We searched herbaria systematically beginning with the first collection cabinet (or last in one case) and continuing through as much of the collection

as possible during the time available. Where possible, we reduced the number of herbaria

folders by searching only area specific folders (e.g., New England folders at Harvard

University). We scanned herbaria sheets for pertinent localities and recorded data directly into

Excel spreadsheets. We searched vertebrate collections at Chicago Academy of Sciences,

Northeastern University, and London Museum of Natural History bird collection by scanning

specimen tags, but searched the Harvard ornithological and London Museum of Natural History

mammal collections by searching accession catalogs. Arrangement of collections first by

taxonomy and second by geography, allowed us to limit searching to appropriate geographic

regions. We searched catalogs to reduce handling of specimens and increase search efficiency.

However, taxonomy within catalogs was often not updated and there was no assurance that

specimens were still present in the collection. Most transfers (i.e., accessions) to other

institutions were noted in catalogs, but in some cases disposal of specimens was not recorded.

We were unable to verify the presence of relevant specimens found in catalogs or corroborate the

taxonomic identification in any collection due to time constraints.

Data Manipulation

All data files were checked for errors and converted to Microsoft Excel 2000 format.

Data were standardized to include the following data fields: institution, taxa, catalog number,

original genus, original species, original subspecies or variety, family, updated genus, updated

species, updated subspecies or variety, common name, state, county, specific locality, park,

proximity to park (1 = within park boundaries, 2 = may be within park boundaries, 3 = in county,

4 = in state), collector, date collected, year collected, sex, age, parts/preparation, remarks, type

status, latitude, longitude, and elevation. We accepted the taxonomic identification of all

specimens without verification; however, we did update taxonomy to conform to current

Integrated Taxonomic Information System (ITIS 2002) standards. Proximity to the nearest national park was determined for each specimen based on the categories noted above using maps (DeLorme 1998, 2000), Mapquest (Mapquest.com, Inc. 2002), and Topozone.com (Maps a la Carte, Inc. 2000), on-line mapping software. Category 1 records must have been determined to have been collected within current park boundaries, even if specimens were collected prior to park establishment. Specimens were assigned category 2 if locality information indicated that records were collected within a town where the park resides in whole or in part, but for which we did not have detailed enough information to classify as being definitively inside or outside park boundaries. Category 2 specimens include all records with general township locality information, but little or no more specific information. Records were assigned category 3 when we were able to determine that records were from within the county the park resides in, but definitely not within park boundaries. For example, a record from Concord, Massachusetts was assigned category 2 for MIMA, but another record from Concord, Massachusetts from the Assabet River was determined not to have been collected within park boundaries and was assigned category 3. Category 4 records include all records from any other county within the state and were assigned to the closest (or only) park in that state. Some records were left unassigned because of insufficient locality data.

Statistical Analyses

We determined if the number of category 1 and 2 records varied with the size or age of the park using linear regression analysis. Data were log transformed and the same analyses performed to determine if the relationships were logarithmic. We used plant data from Acadia and the program EstimateS 6.0b1 (Colwell 2001) to explore the use of species accumulation curves to estimate species diversity.

Results

We received responses to requests for information from 212 specimen collections and 5 natural heritage programs (71% of all queried). We assembled data from 78 collections kept at 52 institutions and the Maine Natural Areas program, 57 directly from collection managers, 11 from web-enabled searches (one collection did not have any relevant specimens), and 10 from manual searches (Table 2). We assembled 31,110 specimen records (30,833 categorized 1-4 by locality) (Table 3) of which 4,745 (15%) are from within park boundaries (category 1) and an additional 4,552 (15%) may be from within park boundaries (category 2), but for which we do not have enough information to determine their exact location. The majority of the specimens, 20,224 specimens (66%), were from within the county (category 3) and the remaining 1,312 (4%) from within the state (category 4). We were unable to categorize by locality 277 specimens because 1) we were unable to identify current locality based on a historic place name, 2) there were discrepancies in the locality data, or 3) they could not be assigned to any one park. Specimens were well distributed between 1890 to 2000 with declines in collecting corresponding roughly with world events such as wars and the Great Depression (Figure 1). In addition, collecting efforts declined post-1970's (Figure 1) suggesting a more quantitative approach to the natural sciences and identification of most vertebrate and plant species endemic to the northeastern U.S..

The number of category 1 and 2 records for each park increased logarithmically with the log size of the park ($n = 14$, $p < 0.0001$, $r^2 = 0.73$) (Figure 2). The number of records did not increase linearly ($n = 14$, $p > 0.05$) (Figure 3) or logarithmically ($n = 14$, $p > 0.2$) (Figure 4) with the age of the park.

More than one-third of all records were from Acadia National Park, the largest and oldest park in this study. Although Acadia National Park had the most category 1 and 2 specimens (4,615), CACO (2,180) had specimens in a greater number of institutions (29) (Table 4). Concentrations of category 1 and 2 specimens could be found in any institution, but most specimens were found in the museums with the largest collections such as the Smithsonian Institution, American Museum of Natural History, Cornell University, and Harvard University (Table 4).

We gathered the most specimens records for plants (13,048) followed by birds (10,056), mammals (5,276), and amphibians and reptiles (2,730). Within the four taxa, we located specimens that totaled 260 families, 909 genera, and 2,055 species/species hybrids (Table 5). Plant specimens were the most diverse taxa with the greatest number of species/species hybrids within the most families and genera (Table 5). Species were further divided among innumerable subspecies and varieties, but these categories were not tallied individually.

Estimates of species richness for plant data for Acaida varied between 850-1100 species (Figure 3) depending on the estimators used.

Discussion

Searching Collections

We assembled a large and diverse group of specimens for 14 parks with minimal staff and limited resources. These data are an important record of biodiversity, providing an excellent baseline dataset from which to evaluate historic biodiversity in our Nation's parks. Despite this wealth of data, however, there are limitations to this information. First, data was not error-checked. We considered each specimen correctly identified and all data accurately recorded. It

is reasonable to expect, however, that either incorrect identification of some specimens or failure to update taxonomy occurred. Nevertheless, we believe the vast majority of identifications to be accurate. Secondly, some data such as locality and date can often be very general or missing altogether. Older specimens often were missing dates or had locality information specific only to a town or county. However, keeping these limitations in mind, the data is extremely valuable as a tool for exploring changes in biodiversity, particularly given the distribution of records over the prior century.

Access to computer records of natural history specimens was critical to this project. Most records received from curators were from collections that had been computerized. Computer records allowed collection managers to complete searches in minutes, rather than days or weeks, that would be necessary to search some collections by hand. Most institutions have some form of computer system that allows retrieval of specimen data by resident curators. In some cases, computerized collections could be accessed by end-users using an internet searching tool, providing easier access to data and reducing time spent by collection managers filling requests for data. These sites often allow the downloading of data to spreadsheet formats and streamlined data gathering. In the future, this work will become increasingly easy to accomplish and bring a wealth of overlooked data to use. Some institutions are years from computerization, not because of technological issues, but because of the time and resources necessary to enter the data. Most natural history collections were unable to search collections themselves because of the lack of time and resources to fulfill such requests, which places the responsibility of searching on the organization requesting the data.

Manual searching of specimen records added 2,517 (8%) records and took 21 days to complete. Although we assembled only 8% of the total records manually, these searches proved

to be an important component of this project and allowed us to access data from specimens that would not have been accessible otherwise In order to retrieve data from these collections manual searches must be conducted. For example, Harvard University Herbaria has less than 1% of their collection computerized of 5,000,000 plant specimens; and within 4 days of searching we searched approximately 2% of the collection and found >800 specimens. Extrapolation would then suggest that there could be as many as 40,000 relevant specimens in that collection alone! However, without manually searching this collection, these data would not be accessible for many years. Searching by hand is time consuming, but can yield valuable data not otherwise available if relying soley only on computer accessible records. In order to be efficient with time, we offer several recommendations. First, knowledge of the localities of interest, including the historical names, is necessary to identify all relevant specimens. Second, lists of potential species for a region can help narrow the search field, with the caveat not to exclude rare, extinct, and vagrant species. Third, we suggest searching specimen tags if division of the collection is by locality. In most large collections, specimens were divided by geographic region into separate folders (plants) or trays (vertebrates). Although the largest collections were daunting in size, they often were the easiest to search because they possessed enough specimens to be divided into small regional categories. Conversely, smaller collections tended to file specimens into locally collected specimens, and then by large geographic divisions (e.g.,North America, outside U.S.) requiring searching of most all of the specimens. Searching specimen tags can be tedious, but has the advantage of updated taxonomy (generally) and the assurance that specimens are still in the collection. However, specimen tags are often very difficult to read, particularly for vertebrates with small tags and old writing. Additionally, handling degrades specimens and may be irritating to the searcher because of the use of harsh chemicals for preservation. Searching by

catalog is much faster, but provides less reliable data and often does not include taxonomic updates. If time permits, searching through catalogs followed by referencing against specimens in the collection may be a good compromise. Management of every collection is different and consideration of those differences is important prior to commencing the search. A flexible strategy is important for determining the best method for conducting searches.

Relationships Between Specimen Distribution And Park Characteristics

It was not surprising that more specimens were found originating from large, natural areas compared to sites established primarily for cultural reasons. Unique natural areas attract collectors of natural history artifacts and generally provide greater opportunity to collect natural history items of interest than cultural sites. It follows then that the two largest parks, Acadia and Cape Cod were responsible for the majority of category 1 and 2 specimens and probably influenced these analyses. However, this trend may also result from the size of the park and thus the area available for collection. In fact, the observed pattern of increasing number of specimens with park size and park type (cultural site versus natural area) is likely a result of both of these factors. In contrast, the lack of a relationship between age of the park and number of records collected probably resulted from the relatively recent establishment of the parks. Specimens were collected throughout the 20th century in many of the areas we now have parks well before they were established, resulting in the lack of relationship with age. Work with specimens originating from parks with a longer established history (i.e., those in the western United States) may provide different results.

Species Diversity Estimators

Estimation of species richness has become an important topic in community ecology and monitoring (Cam et al. 2002) and is an important component of evaluating biodiversity

(Coldwell and Coddington 1994). Species accumulation curves (Soberon and Llorente 1993) have been used to estimate species diversity, but the use of phenomenological models to plot species accumulation data (number of species versus number of sampling occasion or number of specimens collected) have been criticized because there is no mechanistic basis to correct for sampling effort (Fagen and Kareiva 1997, Cam et al 2002). Estimates 6 generates results (Figure 3) using several different estimators (and functions), but the program is vulnerable to the criticisms posed above. Plotting this simple relationship, however, can be useful as an exploratory tool to view the "thoroughness" of sampling conducted in an area and compare sampling across regions (Fagen and Kareiva 1997), or in this case, park units. For parks like Acadia with intensive sampling over several decades, voucher specimens records may provide species richness estimates (Figure 3) that are nearly asymptotic for true species diversity but additional analyses are recommended. Recent statistical procedures such as the information-theoretic approach (Burnham and Andersen 1998) can provide some objectivity in selecting a particular estimator (and function) to determine the accuracy of species accumulation data, assuming a reasonable a priori model set (Cam et al. 2002). Furthermore, other techniques for estimating species richness are preferred to the function fitting approach employed by EstimateS 6 (Cam et al. 2002). The lognormal distribution of species abundances (Fagen and Kareiva 1997), models of detection probability (Cam et al. 2002), and others based on capture-recapture theory are preferred by some authors (see Nichols and Conroy 1996, Boulinier et al. 1998). These models can estimate the size of species assemblages, an important consideration in the design of biological inventories and monitoring programs. We recommend further exploration of how to use these techniques with voucher specimen data in concert with current inventories to provide a comparative measure of linking the past to current conditions.

Literature Cited

Alberch, P. 1993. Museums, collections, and biodiversity inventories. Trends in Ecology and Evolution. 8:372-375.

Bennett, J. P. 2001. Searching for biological specimens from Midwestern parks: pitfalls and solutions. The George Wright forum 18: 26-39.

Boulinier, T., J. D. Nichols, J. R. Sauer, J. E. Hines, and K. H. Pollock. 1998. Estimating species richness: the importance of heterogeneity in species detectability. Ecology 79:1018-1028.

Burgman, M. A., R. C. Grimson, and S. Ferson. 1995. Inferring threat from scientific collections. Conservation Biology 9:923-928.

Burnham, K. P., and D. R. Andersen. 1998. Model Selection and inference, a practical information – theoretic approach. Springer-Verlag.

Cam , E., J. D. Nichols, J. R. Sauer, and J. E. Hines. 2002. On the estimation of species richness based on the accumulation of previously unrecorded species. Ecography 25:102-108.

Colwell, R. K. 2001. EstimateS: Statistical estimation of species richness and shared species from samples. Version 6.0b1. User's Guide and application published at: http://viceroy.eeb.uconn.edu/estimates.

Cotterill, F. P. D. 1995. Systematics, biological knowledge, and environmental conservation. Biodiversity and Conservation 4:183-205.

_____. 1997. The second Alexandrian tragedy, and the fundamental relationship between biological collections and scientific knowledge. Pages 227-241 *in* J. R. Nudds and C. W. Pettit, eds. The value and valuation of natural science collections. Proceedings of the International conference, Manchester, England. Chapman and Hall, London.

DeLorme. 1998. The Maine atlas and gazateer. Yarmouth, ME.

DeLorme. 2000. Topographic Map of Mt. Desert Island & Acadia National Park, Hancock County, Maine. Yarmouth, ME.

Fagan, W. F., and P. M. Karevia. 1997. Using compiled species lists to make biodiversity comparisons

among regions: a test case using Oregon butterflies. Biological Conservation 80: 249-259.

Hawksworth, D. L., and L.A. Mound. 1991. Biodiversity databases: the crucial significance of

collections. Pages 17-29 *in* The biodiversity of microorganisms and invertebrates: its role in

sustainable agriculture. D. L. Hawksworth, ed. CAB International, Wallingford. 302pp.

Integrated Taxonomic Information System (ITIS). 2002. On-line database - http://www.itis.usda.gov.

Huber, J. T. 1998. The importance of voucher specimens, with practical guidelines for preserving

specimens of the major invertebrate phyla for identification. J. Natural History 32:367-385.

McCarthy, M.A. 1998. Identifying declining and threatened species with museum data. Biological

Conservation 83:9-17.

Mapquest.com, inc. 2002. http://www.mapquest.com.

Maps a la Carte, Inc. 2000. http://www.topozone.com.

Nichols, J. D., and M. J. Conroy. 1996. Estimation of species richness. Pages 226-234 in Measuring

and monitoring biological diversity: standard methods for mammals. Wilson, D. E., F. R. Cole,

J. D. Nichols, R. Rudran, and M.S. Foster, eds. Smithsonian Institution Press. Washington, DC.

Ponder, W. F. 1999. Using museum collection data to assist in biodiversity assessment. Pages 253-256

in W. Ponderand D. Lunney, editors. The other 99%: the conservation and biodiversity of

invertebrates. Transactions of the Royal Zoological Society of New South Wales, Mosman.

_____, G. A. Carter, P. Flemons, and R. R. Chapman. 2001. Evaluation of museum collection data for

use in biodiversity assessment. Conservation Biology 15:648-657.

Soberon, J. M., and J. B. Llorente. 1993. The use of species accumulation functions for the prediction of

species richness. Conservation Biology 7:480-488.

Stohlgren, T. J., and J. F. Quinn. 1992. An assessment of biotic inventories in western U.S. National

Parks. Natural Areas Journal 12: 145-154.

U.S. Department of the Interior. 2000. Inventory and Prototype monitoring of

Natural resources In selected national park system units, 1998-1999. National Park Service,

Natural Resource Technical Report NPS/NRI&M/NRTR-2000/1. 58pp.

Table 1. National parks searched for vertebrate and vascular plant voucher specimens.

National Park (Code)	State(s)	Size (Ac)	Year Est.
Northeast Temperate Network			
Acadia National Park (ACAD)	ME	46784	1916
Marsh-Billings-Rockefeller National Historical Park (MABI)	VT	555	1992
Minute Man National Historical Park (MIMA)	MA	967	1959
Morristown National Historical Park (MORR)	NJ	1685	1933
Roosevelt-Vanderbilt National Historic Site (ROVA)[1]	NY	683	1940
Saint-Gaudens National Historic Site (SAGA)	NH	150	1964
Saugus Iron Works National Historic Site (SAIR)	MA	9	1968
Saratoga National Historical Park (SARA)	NY	3406	1938
Weir Farm National Historic Site (WEFA)	CT	60	1990
Coastal and Barrier Island Network			
Assateague Island National Seashore (ASIS)	MD	39732	1965
Cape Cod National Seashore (CACO)	MA	43604	1961
Fire Island National Seashore (FIIS)	NY	19580	1981
Gateway National Recreation Area (GATE)	NY, NJ	26610	1997
Sagamore Hill National Historic Site (SAHI)	NY	83	1963

[1] ROVA was consolidated from Eleanor Roosevelt National Historic Site (ELRO, est. 1977, 181 ac), Home of Franklin D. Roosevelt National Historic Site (HOFR, est. 1945, 290 ac) and, Vanderbilt Mansion National Historic Site (VAMA, est. 1940, 212 ac).

Table 2. Search results for herbaria and vertebrate collections manually searched in this study.

Collection searched	Size (category)[1]	Search time	No. records	Approx. % searched[2]
Herbaria				
Mary Washington College	5,000 (small)	4 hours	1	100
University of South Carolina	85,000 (medium)	3 days[3]	54	100
University of Minnesota	818,000 (large)	4 days	50	17
Harvard University	5,000,000 (large)	4 days	817	2
University of Maryland	67,000 (medium)	2 days	288	33
Vertebrate				
Chicago Acad. Sciences – herp.	20,000 (small)	2 hours	16	25
Northeastern Univ. – vert.	42,000 (medium)	4 hours	148	5
Harvard Univ. – ornith.	338,000 (large)	3 days	504	66
London Mus. Nat. Hist. - birds	2,500,000 (large)	3 days	461	90
London Mus. Nat. Hist. - mammals	359,000 (large)	1 day	178	100

[1] Small ≤30,000, medium ≤195,000), and large >195,000.

[2] Estimated for herbaria based on the number of cabinets searched. Estimates for Chicago depend on the number of rows of specimens searched, for Northeastern are based on the number of cabinet drawers and catalog numbers, and for Harvard are based on catalog numbers. We received help searching from two other people at Chicago and one other at Northeastern.

[3] One day of search time is equivalent to 8 hours.

Table 3. The number of specimen records received in each proximity category for all parks.

Park code	Number of specimen records[1]				Total (%)[2]
	Category 1	Category 2	Category 3	Category 4	
ACAD	3,392	1,223	7,739	149	12,503 (40.6)
MABI	1	199	273	20	493 (1.6)
MIMA	72	408	1,797	78	2,355 (7.6)
MORR	0	119	905	46	1070 (3.5)
ROVA	237	4	251	485	977 (3.2)
SAGA	0	10	102	19	131 (0.4)
SAIR	0	17	722	0	739 (2.4)
SARA	180	6	115	423	724 (2.3)
WEFA	12	15	983	8	1,018 (3.3)
ASIS	471	1	197	3	672 (2.2)
CACO	186	1,994	1,806	6	3,992 (12.9)
FIIS	109	276	4,026	0	4,411 (14.3)
GATE	30	277	1,107	75	1,489 (4.8)
SAHI	55	3	201	0	259 (0.8)
Total (%)	4,745 (15.4)	4,552 (14.8)	20,224 (65.6)	1,312 (4.3)	30,833

[1] Category 1 = within park boundaries, 2 = may be within park boundaries, 3 = in county, 4 = in state.
[2] Totals are reduced by 277 specimens (31,110 total specimens for which we have data),because we were unable to identify current locality based on a historic place name, there were discrepancies in the locality data, or they could not be assigned to any one park.

Table 4. The number of institutions and the top two institutions with category 1 and 2 data for parks in this study.

Park code	No. institutions with cat. 1 and 2 data	Institution with greatest no. cat. 1 and 2 (no.)	Institution with 2nd greatest no. cat. 1 and 2 (no.)
CACO	29	Natural History Museum, London (597)	Harvard University, Museum of Comparative Zoology (401)
ACAD	24	College of the Atlantic (3,110)	University of Maine Herbaria (746)
GATE	20	Smithsonian Institution, National Museum of Natural History (86)	American Museum of Natural History (55)
MIMA	18	Harvard University, Museum of Comparative Zoology (193)	Northeastern University Vertebrate Collection (66)
FIIS	14	Cornell University Museum of Vertebrates (204)	American Museum of Natural History (78)
MORR	9	American Museum of Natural History (48)	Cornell University Museum of Vertebrates (26)
MABI	8	Yale University, Peabody Museum of Natural History (78)	University of California, Berkeley, Museum of Vertebrate Zoology (63)
ASIS	8	University of Maryland, Norton-Brown Herbarium (288)	Smithsonian Institution, National Museum of Natural History (167)
ROVA	5	American Museum of Natural History (160)	New York State Museum (70)
SAIR	5	Harvard University, Museum of Comparative Zoology (9)	Harvard University Herbaria (5)
SARA	5	New York State Museum (157)	Museum of Southwestern Biology (21)
SAHI	4	University of South Carolina, A. C. Moore Herbarium (51)	Natural History Museum, London (4)
WEFA	2	American Museum of Natural History (13)	University of Connecticut, EEB (9)
SAGA	1	American Museum of Natural History (10)	None

Table 5. Taxonomic diversity and total number of category 1 to 4 specimens.

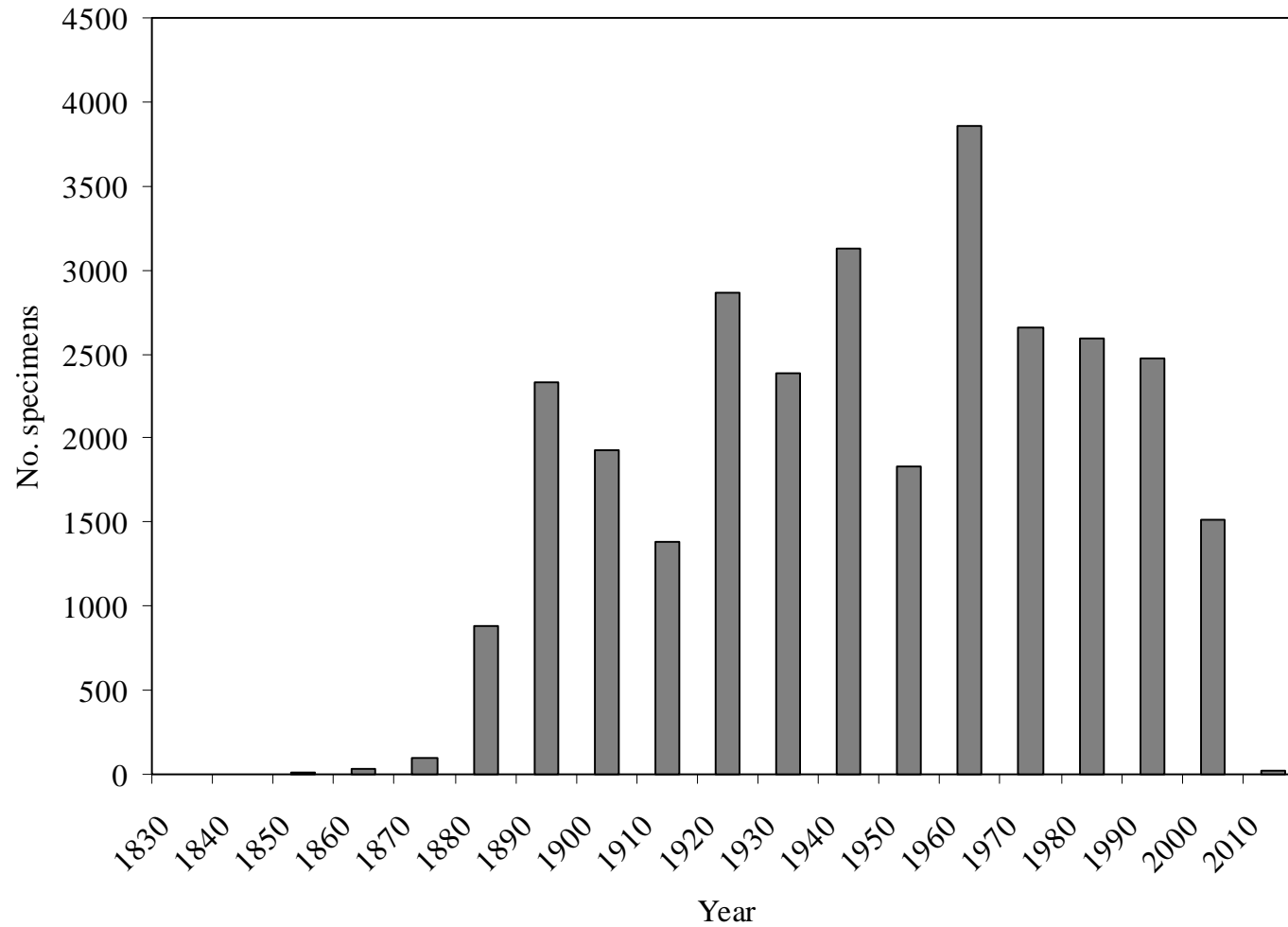| Taxa	No. Specimens	Family	Genera	Spp./hybrids
Birds	10,056	59	196	353
Herps	2,730	18	46	75
Mammals	5,276	32	75	111
Plants	13,048	151	592	1,516
Total	31,110	260	909	2,055

Figure 1. Distribution of year of collection for specimens for the study of vouchers specimens in the northeast. Declines in collecting correspond roughly world events like wars and the Great Depression. A declining trend from the 1970's suggests a more quantitiative approach to the natural sciences and documentation of most flora and fauna species in the northeastern U.S.
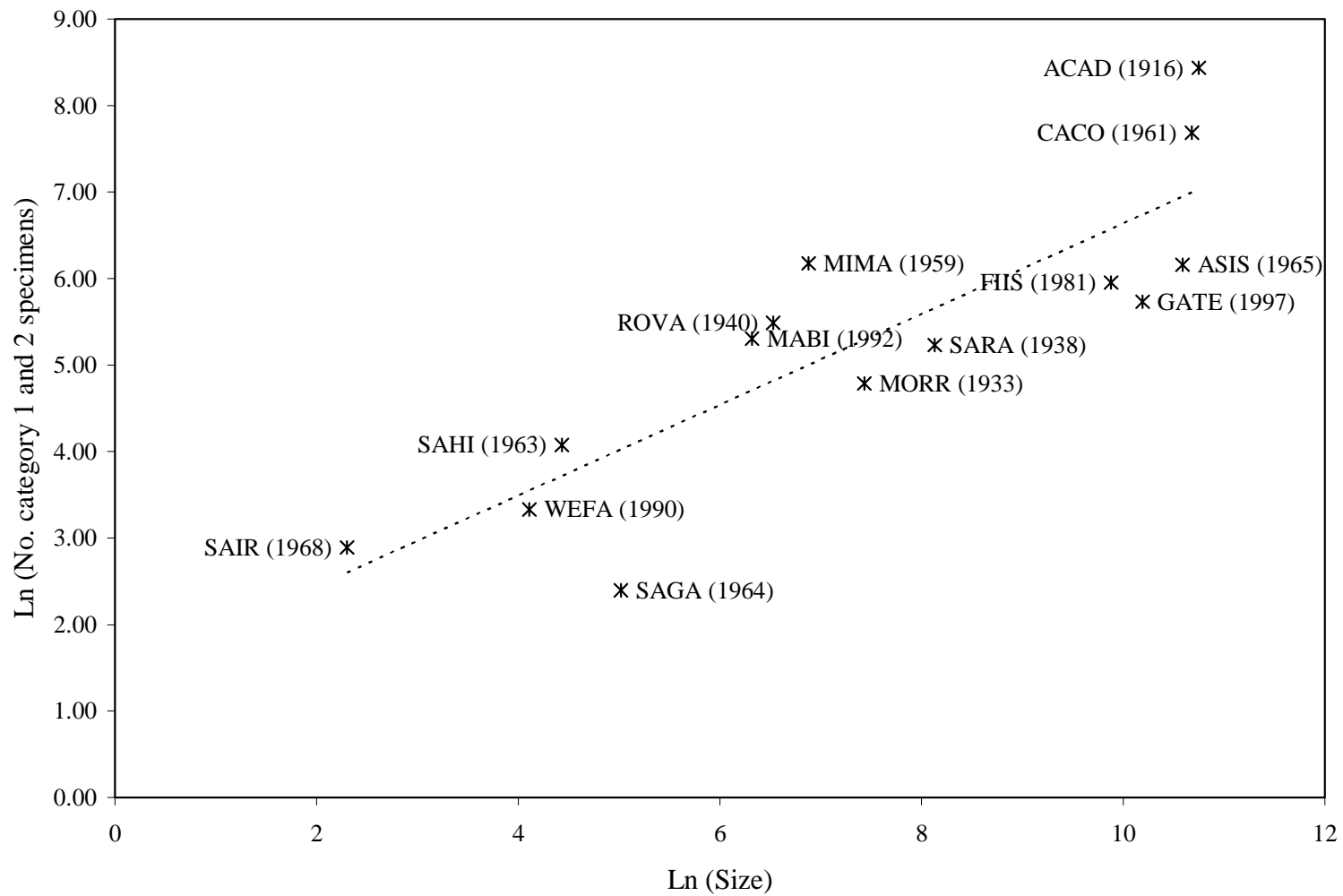
Figure 2. The natural log of the number of category 1 and 2 specimens vs. the natural log size of the park (in acres) in which they are associated. Parks are identified by park codes and includes the year established.
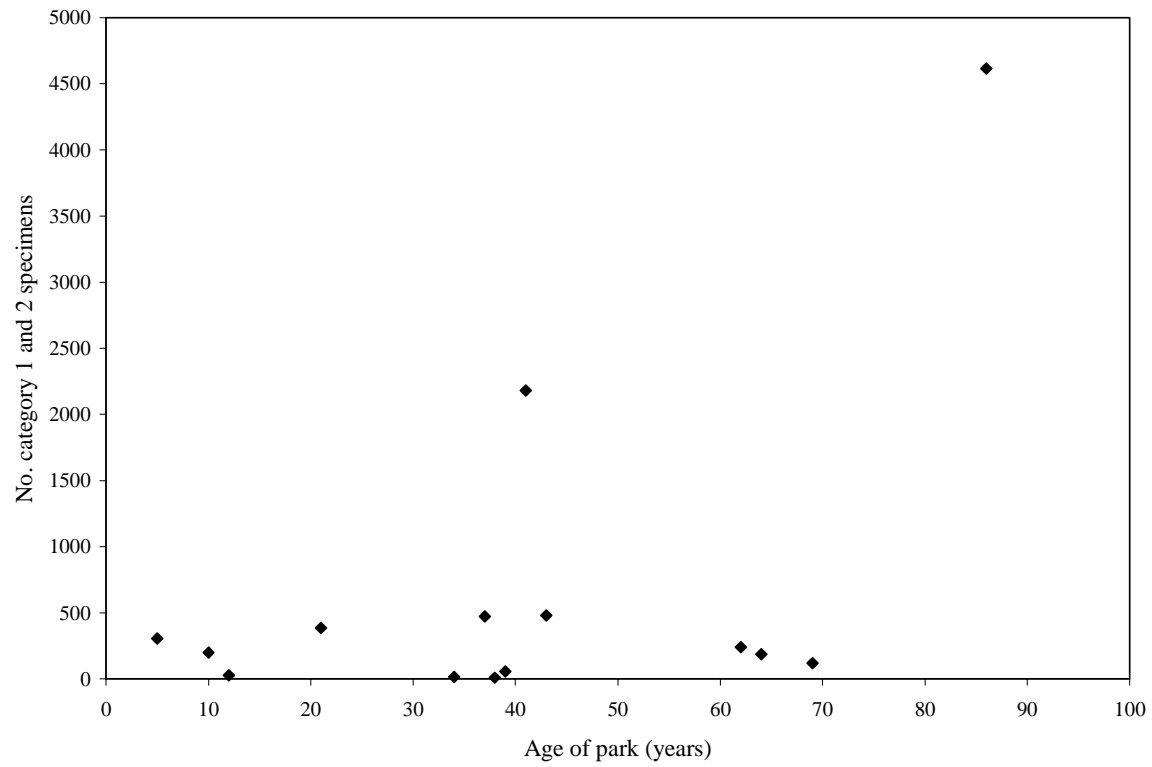
Figure 3. The number of category 1 and 2 specimens vs. the age of the park in which they are associated.
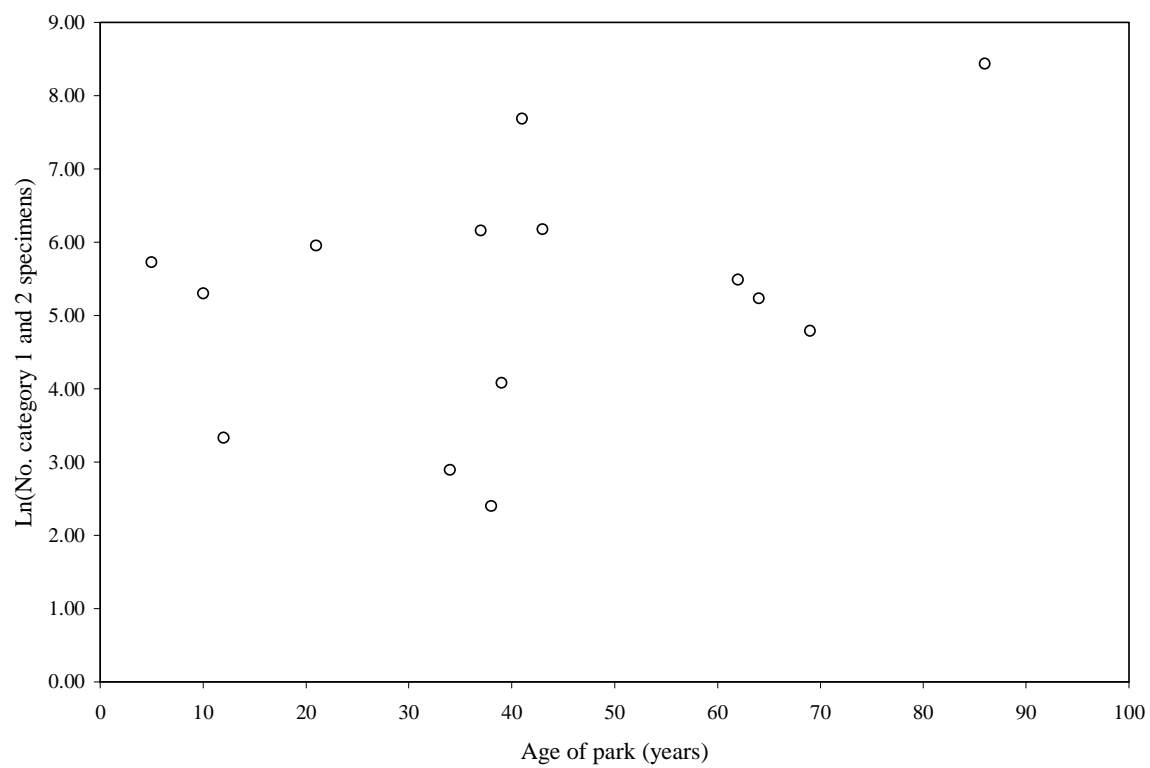
Figure 4. The natural log of the number of category 1 and 2 specimens vs. the age of the park in which they are associated.

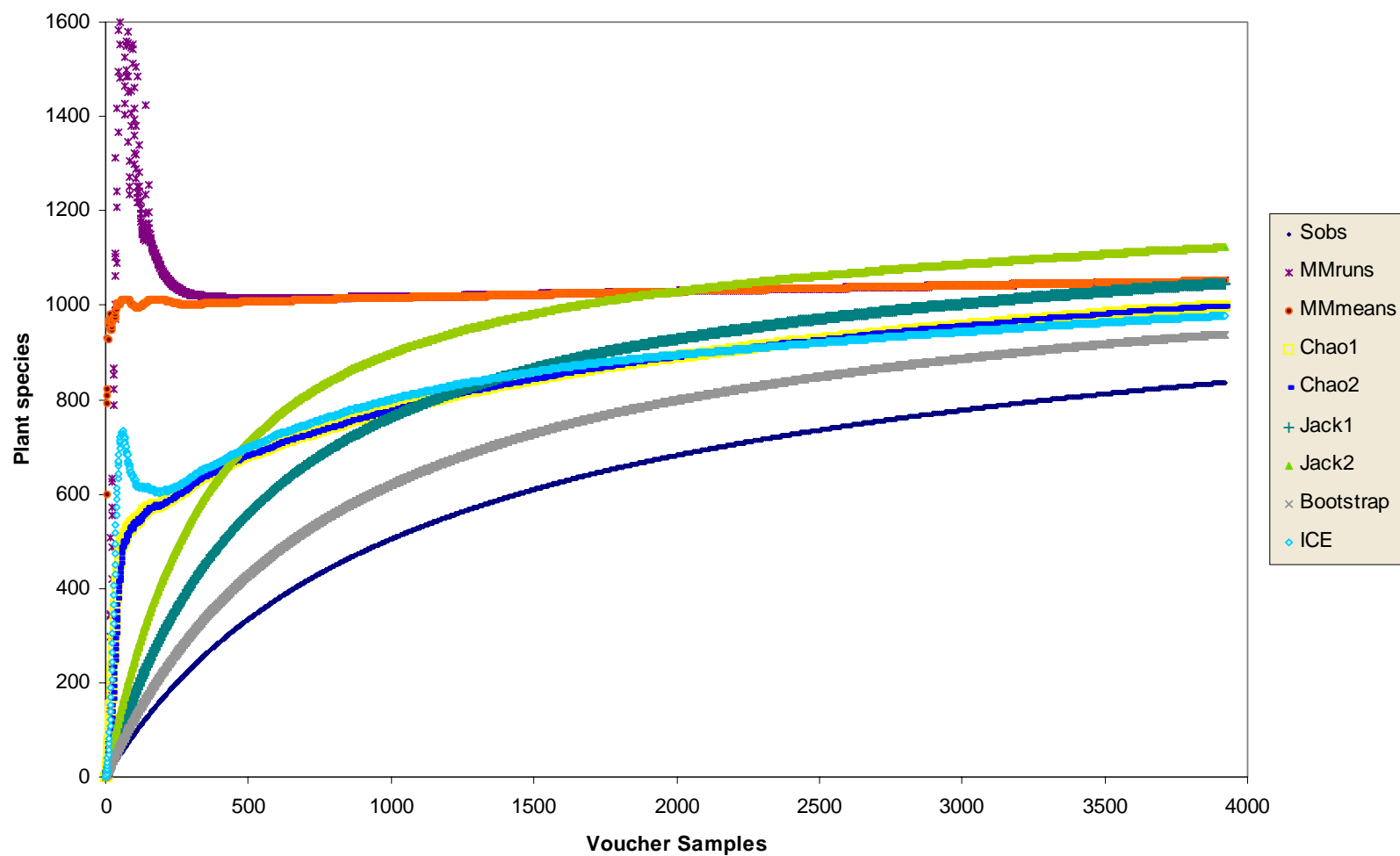**ACAD Plant diversity by adding plant voucher specimens**

Figure 3. Plant species diversity for Acadia National Park plotted against the number of voucher specimens identified (category 1 & 2) using the program EstimateS 6. Results are based on actual observations and numerical estimators (see Discussion Section).